行政院國家科學委員會專題研究計畫 期末報告

估計機率分佈演算法求解群內最佳化問題之研究

計 畫 類 別 ： 個別型
計 畫 編 號 ： NSC 101-2221-E-230-026-
執 行 期 間 ： 101 年 08 月 01 日至 102 年 10 月 31 日
執 行 單 位 ： 正修科技大學資訊管理系

計 畫 主 持 人 ： 陳世興

計畫參與人員： 碩士班研究生-兼任助理人員：洪毅
碩士班研究生-兼任助理人員：李易霖
碩士班研究生-兼任助理人員：吳奇燊
碩士班研究生-兼任助理人員：林文裕
大專生-兼任助理人員：曾昶瑜
大專生-兼任助理人員：張瑋庭

報 告 附 件 ： 出席國際會議研究心得報告及發表論文

公 開 資 訊 ： 本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中　華　民　國　102 年 10 月 21 日

中文摘要： 估計機率分佈演算法(Estimation of Distribution Algorithms, EDAs)在最近已成為演化式演算法中重要領域之一，可用來解決困難之組合性最佳化問題，但本計畫主持人發現過去的研究僅有一個可求解群內最佳化問題之 EDAs，如多旅行推銷員(Multiple Traveling Salesmen Problems, mTSP)或平行機台排程問題(Parallel Machine Scheduling Problems, PMSPs)，這些問題同時都包含了指派與順序之最佳化。基於過去少有 EDAs 演算法求解群內最佳化問題之相關研究，本研究提出一個自指引基因演算法並結合最小負載指派法則 (Minimum Loading Assignment rule, MLA rule)，用此轉化的編碼方式 (Transform-Based Encoding)來切入 mTSP 這類型的問題，其求解間僅 n!，此結果與目前最佳之直接編碼方式兩段式染色體編碼遺傳演算法(Two-Part Encoding Genetic Algorithm, TPGA) 比較，兩段式染色體編碼求解間仍需 n!*C(n-1, m-1)，因此預期所提出方法會比 TPGA 佳。
所採用的三十三個測試例題從 TSPLIB 取得，目標函數包含了極小化總移動距離與極小化最大移動距離，推銷員人數有 2、3、5、10 與 20 人，實驗的結果顯示本研究所提出的 SGGA 演算法結合 MLA 法則後，無論在極小化總移動距離與極小化最大移動距離，都能比目前最佳的直接編碼方式佳。且在極小化總移動距離的問題中，在使用三個至十個推銷員時，目標函數值並不會隨著人數增加而隨之成長，因此本研究建議之後的 EDAs 研究，採用轉化的編碼方式會比直接編碼佳。

中文關鍵詞： 估計分佈演算法、群內最佳化問題、多旅行推銷員問題、等效平型機台排程問題

英文摘要： Estimation of Distribution Algorithms (EDAs) have recently been recognized as a prominent alternative to traditional evolutionary algorithms due to their increasing popularity. The core of EDAs is a probabilistic model which directly impacts performance of the algorithm. Previous applications of EDAs have used algorithms to solve many hard problems. However, this researcher found that there is one problem which EDAs does not discusses so far. It is the in-group optimization problems, such as the multiple traveling salesmen problem (mTSP) and parallel machine scheduling problem (PMSP) studied in this research. These problems include the assignment and sequencing procedures in the same time and to be shown in different forms. As a result, this research proposed an algorithm deal by using the Self-Guided GA together with the Minimum Loading Assignment rule (MLA) to tackle the mTSP. This strategy is called the transformed-based encoding approach instead of the direct encoding. The solution space of the proposed method would be only n!. We compare the proposed algorithm against the best direct encoding technique, two-part encoding genetic algorithm (TPGA) and its solution space is n!*C(n-1, m-1), in the experiment on the 33 instances drawn from the well-known TSPLIB. The experimental results show the proposed algorithm is better than the two-part encoding genetic algorithm in terms of minimization of the total traveling distance and the maximum traveling distance among the salesmen. An interesting result also presents the proposed algorithm would not cause longer traveling distance when we increase the number of salesmen from 3 to 10 persons under the objective of minimization of total traveling distance. Consequently, this research may suggest the EDAs researcher could employ the MLA rule instead of the direct encoding in their proposed algorithms.

英文關鍵詞： Estimation of Distribution Algorithms, In-Group Optimization Problems, Multiple Traveling Salesmen Problems, Identical Parallel Machine Scheduling Problems.

2 緣由與目的：

Estimation of Distribution Algorithms (EDAs) uses the learning while optimizing principle [16] and have emerged as a prominent alternative to evolutionary algorithms [7, 10, 32, 13, 21, 26]. Compared with genetic algorithms (GAs) that employ the crossover and mutation operators to generate solutions, EDAs do not use the crossover or mutation. Instead, they explicitly extract global statistical information from the previous search and build a posterior probability model of promising solutions from which new solutions are sampled. It is the most important characteristic to distinguish EDAs from GAs [31].

A number of the latest papers on EDAs in solving some NP-hard scheduling problems [14, 7, 5, 9, 32, 21] have shown that EDAs are able to perform effectively. Ceberio et al. [5], in particular, extensively tested 13 famous permutation-based approaches in EDAs on four well-known combinatorial optimization problems, including Travelling Salesman Problem (TSP), Permutation Flowshop Scheduling Problems (PFSPs), Linear Ordering Problem (LOP), and Quadratic Assignment Problem (QAP). Their paper has provided a good basis for comparison.

Even though EDAs was effective in solving various hard problems, there is a problem that EDA is not discussed extensively. To the best of our knowledge, only one EDAs proposed by Shim et al. [26] is able to solve in-group optimization problems, such as the Multiple Traveling Salesmen Problems (mTSP) and the Parallel Machine Scheduling Problems (PMSPs) belonged to this category [2]. In-group optimization problems involve the assignment and routing/sequencing procedures in the same time. Take mTSP for example, a number of n cities are assigned to m salesmen and these n cities are visited once by a salesman where $n > m$. It is apparently that this problem is a NP-Hard problem.

Due to there was only a few EDAs could solve the in-group optimization problems, there is much room to do the research on them. Besides, in-group optimization problems are very practical in industry, such as the application of mTSP. this research proposed an algorithm deal by using the Self-Guided GA together with the Minimum Loading Assignment rule (MLA) to tackle the mTSP. This strategy is called the transformed-based encoding approach instead of the direct encoding. The solution space of the MLA would be only $n!$. We compare the proposed algorithm against the best direct encoding technique, two-part encoding genetic algorithm (TPGA)[4], in the experimental section. It is notable that solution space of the two-part encoding approach is $n!\binom{n-1}{m-1}$. The proposed method MLA, consequently, is better than the two-part encoding technique. A better solution quality is expected when SGGA works with MLA method.

3 研究報告應含的內容

3.1 Encoding Methods for solving in-group optimization problems

This project already advocates the importance of the research direction in previous sections. Because the exact algorithms [12, 18, 20] may not tackle with the large size problems efficiently, evolutionary algorithms (EAs) is one of the commonly used. The first important step of using EAs is to select the appropriate encodings. The solution representations may fall into the two classes, the direct and the transformed-based encoding methods [2].

There are five major direct encodings of the EAs, including the one-chromosome [28], two-chromosome [19, 22], two-part chromosome [4], matrix representation [1] and grouping genetic algorithms [3, 11, 17, 23, 27]. Then, [1] proposed a matrix representation of the N jobs on M machines, whose size is $M \times N$. Each row indicates that the parallel machines and the processing sequence of the jobs on it. When there is no jobs to be processed on that machine, number 0 is inserted to the blank spaces. As a result, it is apparently that the memory usage is not efficient. That is, $M \times N - N$ spaces are unused if we apply this encoding technique. In GGA, it commonly uses an array of jobs for for each machines which shows the processing order of the jobs assigned to that machine [29]. Kashan et al. [15] further extends the GGA into the grouping version of the particle swarm optimization algorithm.

In these direct encoding techniques, the best approach could be the two-part chromosome technique according to Carter and Ragsdale[4]. When we have n items and m groups, the solution space of one chromosome needs $(n + m - 1)!$. Two chromosome approach takes $n!m^n$ and the size of the two-part chromosome is $n!\binom{n-1}{m-1}$. Due to the two-part chromosome technique takes less solution space to do the explorations, this study selects this encoding technique to be compared with the proposed EDAs.

The first estimation of distribution algorithm for mTSPs is to apply the one-chromosome representation [26]. Because there are $m-1$ pseudo cities introduced in the chromosome, every chromosome consists of $n+m-1$ genes. As a result, the dimension of their probability model $P_r(x)$ by computing the marginal probability of each city is $N \times N$ where N is $n+m-1$.

The second major encoding type is to separate sequencing and assignment decisions because the complex encoding may yield poor results [24]. Its encoding strategy utilizes the permutation encoding first and then assign the items onto the groups at every stage. This separated method is applicable on the complex flowshop problems when there are many parallel machines in the flowshop. Ruiz and Maroto[24] names it as the priority rules for hybrid flowshops. Wang et al. [30] calls it as the earliest completion factory (ECF) rule for solving the distributed permutation flow-shop scheduling problem. Salhi et al. [25] selects the index of the machine that allows a job which has the fastest completion time for solving the complex flowshop scheduling problems. Because the transformed-based encoding method might be efficient, this research adopts this approach instead of the direct encoding. In addition, several EAs could apply the assignment rule and then solve the in-group optimization problems. To evaluate the performance of the algorithms studied in this research, we select the mTSP to do the extensive comparisons.

3.1.1 Assignment Rule in the mTSP problems

Given a set of city sequence $\pi_1, \pi_2, ..., \pi_n$ in π and these cities are not assigned to any salesman yet. This sequence π could be decoded to by assigning the cities to salesmen. That is, the this assignment rule is executed in the fitness function of each chromosome. The rule we called is the minimum loading assignment (MLA) rule. The following pseudo code illustrates the MLA rule.

Algorithm 1 Minimum loading assignment rule

Require:

i: The position of a city in the sequence π

$k[i]$: The current number of assigned cities of a salesman i

$\Omega^i_{k[i]}$: The

$i \leftarrow 1$
while $i \leq m$ **do**
  $k[i] \leftarrow 1$
  $\Omega^i_{k[i]} \leftarrow \pi_i$
  $i \leftarrow i+1$
  $k[i] \leftarrow k[i]+1$
end while
while $i \leq m$ **do**
  Select a salesman j who could process the π_i with the minimum objective value
  $\Omega^j_{k[j]} \leftarrow \pi_i$
  $i \leftarrow i+1$
  $k[i] \leftarrow k[i]+1$
end while

In the beginning, the first m cities are assigned to the m salesmen and we calculate the objective values of each salesman. The objective function of mTSP would be the total traveling distance or the maximum traveling distance among the salesman. After that, we do the MLA rule iteratively for the unassigned cities. MLA rule assigns the first unassigned city in the sequence π to a salesman when it causes the minimum objective value. This assigned city is removed from the π. This rule is not stopped until there is no city in the π. By using the rule, it means the assigned city could be assigned to a salesman who has the less loading. It also implies that this assigned city might be closed to the last city visited by the salesman so that a far away city would not be considered. Through the MLA rule, it is able to be extended to the parallel machine scheduling problem with setup consideration or the distributed flowshop scheduling problem.

3.1.2 Transformed-Based Encoding in Self-Guided Genetic Algorithm

After we introduced the assignment rule in mTSP, this section describes the detail procedures of the Self-guided GA. The benefits of the proposed method are preserving the salient genes of the chromosomes, and exploring and exploiting good searching directions for genetic operators. In addition, since the probabilistic difference provides good neighborhood information, it can serve as a fitness function surrogate. The detailed procedure of the Self-guided GA is described as follows:

Step 1 is the initialization of a population. The sequence of each chromosome is generated randomly.

Step 2 initializes the probability matrix $P(t)$ and the matrix size is $n-by-n$, where n is the problem size. Step 7 builds the probabilistic model $P(t)$ after the selection procedure.In Step 8 and Step 9, $P(t)$ is

Algorithm 2 MainProcedure of Self-guided GA()

Population: A set of solutions
Generations: The maximum number of generations
$P(t)$: Probabilistic model
t : Generation index

```
Initialize Population
t ← 0
Initialize P(t)
while t < generations do
    EvaluateFitness (Population)
    Selection/Elitism(Population)
    P(t + 1) ← BuildingProbability-Model(Selected Chromosomes)
    Self-Guided Crossover()
    Self-Guided Mutation()
    t ← t + 1
end while
```

employed in the self-guided crossover operator and the self-guided mutation operator. The probabilistic model will guide the evolution direction, which is shown in Section 3.1.4 and Section 3.1.3. In this research, the two-point central crossover and swap mutation are applied in the crossover and mutation procedures for solving the mTSP under this study.

We explain the proposed algorithm in detail in the following sections. We explain how the probabilistic model guides the crossover and mutation operators.

3.1.3 Crossover Operator with Probabilistic Model

The idea of Self-Guided Crossover is the same with Self-Guided Mutation, which employs the probability differences of the mating chromosomes by using the Eq. 1. By doing so, we could evaluate which chromosome is mated with a parent solution. For the detail description, please refer in [6].

$$\Delta = \Delta_1 - \Delta_2 = \prod_{p \in (CP1\ to\ CP2), g=[p]}^{n} P(Candidate1_{gp}) - \prod_{p \in (CP1\ to\ CP2), g=[p]}^{n} P(Candidate2_{gp}). \quad (1)$$

3.1.4 Mutation Operator with Probabilistic Model

Suppose two jobs i and j are randomly selected and they are located in position a and position b, respectively. p_{ia} and p_{jb} denote job i in position a and job j in position b. After these two jobs are swapped, the new probabilities of the two jobs become p_{ib} and p_{ja}. The probability difference Δ_{ij} is calculated as Eq. 2, which is a partial evaluation of the probability difference because the probability sum of the other jobs remains the same.

$$\begin{aligned}\Delta_{ij} &= P(X') - P(X) \\ &\approx \prod_{p \notin (aorb), g=[p]}^{n} P_{t+1}(X_{gp})[(p_{ib}p_{ja}) - (p_{ia}p_{jb})].\end{aligned} \quad (2)$$

Now that the part of $\prod_{p \notin (aorb), g=[p]}^{n} P_{t+1}(X_{gp})$ is always ≥ 0, it can be subtracted and Eq. 2 is simplified as follows:

$$\Delta_{ij} = (p_{ib}p_{ja}) - (p_{ia}p_{jb}). \quad (3)$$

$$\Delta_{ij} = (p_{ib} + p_{ja}) - (p_{ia} + p_{jb}). \quad (4)$$

If Δ_{ij} is positive, it implies that one gene or both genes might move to a promising area. On the other hand, when Δ_{ij} is negative, the implication is that at least one gene moves to an inferior position.

On the basis of the probabilistic differences, it is natural to consider different choices of swapping points during the mutation procedure. A parameter TM is introduced for the self-guided mutation operator, which denotes the number of tournaments in comparing the probability differences among the TM choices in swap mutation. Basically, $TM \geq 2$ while $TM = 1$ implies that the mutation operator mutates the genes directly without comparing the probability differences among the different TM choices.

When $TM = 2$, suppose the other alternative is that two jobs m and n are located in position c and position d, respectively. The probability difference of exchanging jobs m and n is:

$$\Delta_{mn} = (p_{md} + p_{nc}) - (p_{mc} + p_{nd}). \quad (5)$$

After Δ_{ij} and Δ_{mn} are obtained, the difference between the two alternatives is as follows:

$$\Delta = \Delta_{ij} - \Delta_{mn}. \quad (6)$$

If $\Delta < 0$, the contribution of swapping job m and n is better, so we swap job m and n. Otherwise, jobs i and j are swapped. Consequently, the option of a larger probability difference is selected and the corresponding two jobs are swapped. By observing the probability difference Δ, the self-guided mutation operator exploits the solution space to enhance the solution quality and prevent destroying some dominant genes in a chromosome. Moreover, the main procedure of the self-guided mutation is Eq. 6, where the time-complexity is only a constant after the probabilistic model is employed. This approach proves to work efficiently.

To conclude, the Self-guided GA is obviously different from the previous EDAs. Firstly, the algorithm utilizes the transformed-based encoding instead of using the direct encoding used by Shim et al. [26]. Secondly, the proposed algorithm explicitly samples new solutions without using the crossover and mutation operators. The Self-guided GA embeds the probabilistic model in the crossover and mutation operators to explore and exploit the solution space. Most important of all, the algorithm works more efficiently than previous EDAs [26] in solving the mTSP because the time-complexity is $O(n)$ whereas the previous EDAs needs $O(n^2)$ time.

3.2 Experimental Results of the Proposed Algorithm

We conducted extensive computational experiments to evaluate the performance of Self-guided GA together with the MLA rule in solving the mTSP. The proposed algorithm was compared with the benchmark encoding algorithm, Two-Part chromosome GA, from the literature [4]. In addition, we employ the genetic operators and parameter settings of Two-Part chromosome genetic algorithm suggested Chen and Chen [8]. The genetic operators are the two-point crossover operator and the swap mutation operator. As a result, it ensures we do a fair comparison between the proposed algorithm with the benchmark encoding algorithm. Besides, a standard genetic algorithm (SGA) also applies the MLA rule which could show the performance enhanced by the assignment rule proposed by this research.

The objective functions include minimize the total traveling distance and the maximum traveling distance which are shown in Section 3.2.1 and Section 3.2.2, respectively. We implemented the algorithms in Java 2 on a Windows 2003 server (Intel Xeon 3.2 GHZ).

Across all the experiments, we replicated each instance 30 times on the 33 instances from the well-known TSPLIB. We assume the first city of each instance is the home-depot. The size of these instances is from 48 to 400. The number of salesmen is ranging from 2, 3, 5, 10, and 20. As a result, we conduct extensive experiments to evaluate the proposed algorithm under different circumstances.

3.2.1 Results of the total traveling distance

The firsrt objective evaluates the total distances travelled by the m salesmen. It reflects the total cost of the assignment. Fig. 1 shows the main effects plot on the method comparison and the differences of the number of salesmen we assign. This figure clearly illustrates the SGGA and SGA (denoted $GA_Heuristic$) are better than the Two-Part encoding GA (named $GA_TwoPart$). It means the MLA rule, i.e. the transformed-based method, could be a promising approach which is better than the direct encoding method. Then, when the number of salesmen increased, especially there are 20 salesmen could be assigned, the total distance is increased greatly. As a result, it implies the inefficiency if we request too many salesmen in terms of the managerial perspective.

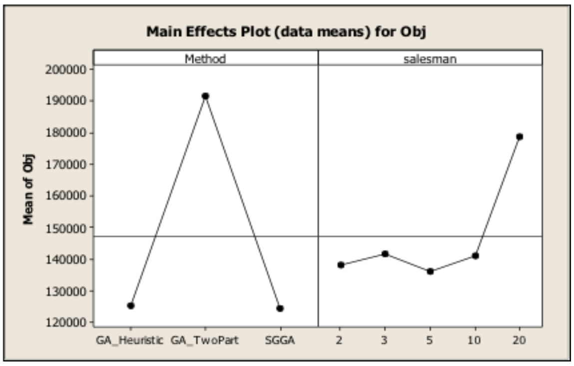

Figure 1: Main effects plot on the total traveling distance of the compared algorithms

Fig. 2 depicts the interaction plot between the factor method and the number of salesmen. It might be interesting to see the SGGA and SGA that do not yield the longer total traveling distance when the number of salesmen increased from two to 10 salesmen. However, Two-Part encoding GA may auffer the pain of the number of salesmen increased. This figure could distinguish the effectiveness for the transform-based rule to the direct encoding method.

Finally, if a manager would like to determine how many salesmen is required, the lowest total traveling distance would be ten according to this interaction plot.

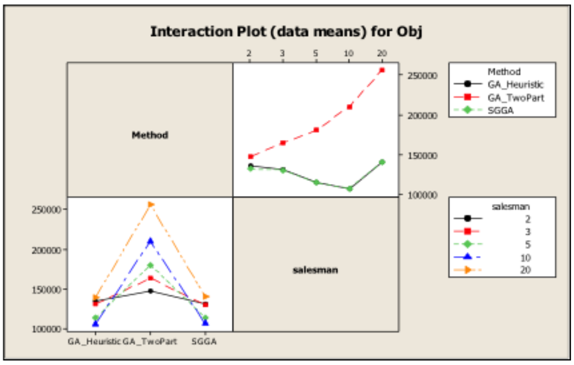

Figure 2: Intreaction plot on the total traveling distance of the compared algorithms

Finally, the detail result of the three compared algorithms is shown in Table 1.

3.2.2 Results of the maximum traveling distance

When it comes to the second objective tested by the three algorithms, the maximum traveling distance is used. The algortihms would minimize the loading of a salesman who has the highest loading. As a result, this objective presents the balanced loading among the salesmen. In Figure 3, both SGGA and SGA are remain better than Two-Part encoding GA while SGGA is slightly better than SGA. The primary reason would be the MLA rule which selects a suitable salesman during the assignment phase. The maximum traveling distance, thus, is reduced based on the rule.

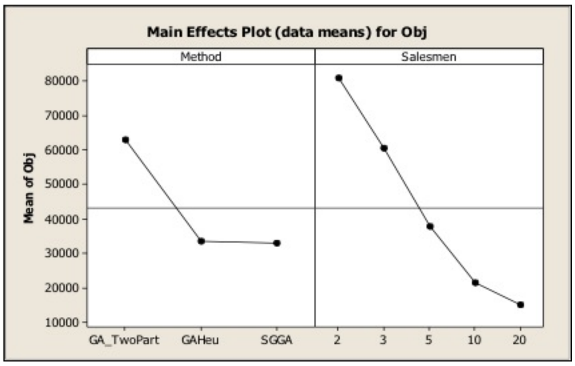

Figure 3: Main effects plot on the maximum traveling distance of the compared algorithms

When there are tweenty salesmen are assigned (see Fig. 4, it causes the lowerest maximum loading of a salesman. It is a reasonable result because the loading could be distributed to many salesmen. Compared the results obtained by prior section, if a manager assigns 20 salesmen, it causes the longest total traveling distance. Hence, the two objectives present the trade-off and should be considered simultaneously.

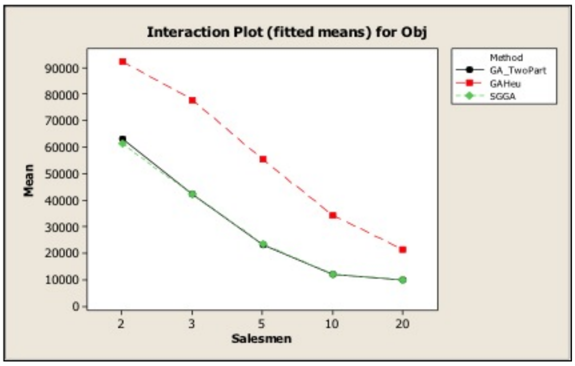

Figure 4: Intreaction plot on the maximum distance of the compared algorithms

Table 2 is the complete results of the three compared algorithms. It would show the SGGA and SGA is better than the Two-Part encoding GA.

4 Conlcusions

This study solve the in-group optimization problems which is rarely solved by the EDAs. A new EDAs SGGA was proposed, which works with the MLA rule together. In addition, because the MLA rule is classified in the category of transform-based encoding, the proposed algorithm is compared with the two-part encoding GA which is is the best direct encoding strategy so far. We evaluate these algorithm by solving the mTSP problem under 33 instances drawn from TSPLIB. The experimental results show the SGGA with the MLA rule outperforms the Two-Part encoding GA in both the total traveling distance and the maximum traveling objectives. It reveals the proposed algorithm is capable for solving the mTSP problem well. In addition, the MLA rule is also effective and could be applied on some GAs that originally designed for the permutation type problems. As a result, this research provides an insightful results for the researchers who are doing the scheduling problems and could move toward the in-group oprimization problems.

Table 1: Total Distance

	$GA_Heuristic$			$GA_TwoPart$			SGGA		
instance	Minimum	Mean	Maximum	Minimum	Mean	Maximum	Minimum	Mean	Maximum
att48	39873	67412	119821	41382	80572	142234	38279	67408	121381
berlin52	9194	13307	20647	9911	15078	23586	8773	13397	21140
bier127	220310	257767	329280	238090	318724	433197	214928	257983	311903
ch130	12781	16173	19262	15896	22951	32313	12020	15959	19035
eil101	1024.4	1251.1	1625.4	1170.4	1734.2	2441.7	910.5	1230.2	1624.1
eil51	497.2	726.6	1195.1	566	896.7	1450.8	483.7	726.2	1186
eli76	771.2	1036.2	1640.5	840.8	1365.6	2151.2	697.2	1030.3	1610.3
gr96	911.3	1325.6	2260.3	1103.9	1834.4	2968.2	893	1325.3	2257.7
kroa150	62801	80353	104886	84155	118336	164472	60914	77752	93631
kroa200	90234	114171	145046	122311	164758	223643	86536	112027	141175
kroB100	39466	52960	74861	49619	78483	118140	39213	52689	78913
kroB150	62532	80549	93700	81609	122203	171691	63013	79185	97548
kroB200	85875	114978	149151	117037	162784	220721	88753	113813	142722
kroC100	37895	53097	77495	55150	81016	128239	37420	51764	78388
kroD100	40864	54795	82342	51506	80103	129610	38858	53964	81408
kroE100	39623	55741	80784	52077	82172	128825	38145	55008	85800
lin105	30255	45563	79901	37418	65266	107157	26922	45533	77804
lin318	182437	243339	322419	251134	328117	447096	183700	243829	318757
mtsp100	40596	54269	81858	50939	80380	123128	39229	54107	83089
mtsp150	76189	96634	116498	101690	138019	184927	76317	94583	118046
mtsp51	467.2	725.2	1188	558.9	897.7	1436.8	475.1	727.6	1198.2
pr124	133373	201905	298646	199047	328870	507277	134543	198651	296597
pr136	223163	282285	383137	273500	419743	634173	200711	281343	390805
pr144	194333	253771	327206	265350	399668	587282	198892	251187	323671
pr152	207562	309617	409260	332865	524397	790537	217451	304029	403995
pr226	329781	519456	753744	588036	842705	1223229	343841	517697	728401
pr264	172047	244558	386775	360147	504250	698832	173869	239184	314209
pr299	218292	283892	361192	314827	413047	585709	220328	282296	357225
pr76	167847	251580	434977	185337	334205	560341	148840	252445	440088
rat195	7200.4	8775.9	10212.1	8563	12055	16873	7059.3	8667.7	10285.1
rat99	2144.1	3418.9	5939.7	2768	4606	7707	2013.9	3392.5	5924
rd400	53411	87481	126827	91499	111470	135460	53204	88769	122815
st70	939.3	1497.3	2594	1142.5	2005.8	3315.9	929.4	1495.7	2641.4
tsp225	13615	16933	20478	17201	23387	32360	13417	16975	20124

Table 2: Maximum Distance

	$GA_TwoPart$			GAHeu			SGGA		
instance	Minimum	Mean	Maximum	Minimum	Mean	Maximum	Minimum	Mean	Maximum
att48	13678	20860	35275	13668	16292	25270	13668	16345	26080
berlin52	2440	4237	7409	2440	3340.9	6092.5	2440	3352.8	6295.9
bier127	34697	106128	196001	24007	68454	157367	24007	67228	150031
ch130	2472	7449	13013	1177	4277	9982	1177	4214	9967
eil101	178	525.1	945	107.2	312.9	731	108	311.1	733.2
eil51	109	228.72	443	108.89	168.21	336.21	108.89	166.35	319.37
eli76	141	374.3	702	124.3	242	528.4	124.4	241.5	509.9
gr96	204	504.9	912	169.7	311.9	632.9	169.7	313.5	644.4
kroa150	12890	40794	76751	5395	21450	51331	5395	20930	51474
kroa200	18234	56814	105496	6221	30678	73756	6271	30173	74055
kroB100	8623	23496	44141	6699	13432	28484	6699	13287	29348
kroB150	13835	39730	74536	5750	20810	51433	5750	20353	45182
kroB200	17701	55980	102982	6698	30614	74246	6697	30487	70855
kroC100	8456	24044	43517	5750	12989	27426	5750	12590	26317
kroD100	8676	23573	41560	6357	13350	28560	6357	13199	28950
kroE100	8896	24369	43138	7038	13838	29229	7038	13835	27412
lin105	7334	18067	31798	6375	10973	21901	6375	10770	22956
lin318	34679	113340	211238	10175	65604	161115	10266	66258	159562
mtsp100	8670	23730	43101	6357	13589	29952	6357	13287	30031
mtsp150	14710	46671	89671	5352	25518	60870	5306	24969	60878
mtsp51	109	229.57	428	108.89	165.75	321.79	108.89	167.32	337.88
pr124	38626	103277	187550	22594	52018	118518	22594	50987	130491
pr136	47037	130878	236763	25731	71766	163532	25731	70236	158871
pr144	46297	127255	226373	24313	68093	165328	24313	66713	162475
pr152	59013	165696	301989	31727	82679	197520	31727	83184	216330
pr226	91773	291805	550302	34845	147311	387354	34845	145246	370270
pr264	57515	178142	323401	16339	68682	189554	16524	67060	172857
pr299	42242	138823	254915	14293	74439	182661	14344	74295	179707
pr76	38692	88456	162369	37971	58398	113437	37971	58522	114227
rat195	1271	3851	7103	604	2194	5246	605	2157	4798
rat99	503	1238.9	2079	432.4	769.3	1609.4	432.9	763.4	1485.8
rd400	12312	41808	82410	2970	24824	62040	2903	24918	61631
st70	221	516.5	1014	206.4	336.4	705.1	207.4	337.9	701.6
tsp225	2521	7521	13717	998	4348	10209	998	4354	9966

References

[1] J.-P. Arnaout, G. Rabadi, and R. Musa, "A two-stage ant colony optimization algorithm to minimize the makespan on unrelated parallel machines with sequence-dependent setup times," *Journal of Intelligent Manufacturing*, vol. 21, no. 6, pp. 693–701, 2010.

[2] T. Bektas, "The multiple traveling salesman problem: an overview of formulations and solution procedures," *Omega*, vol. 34, no. 3, pp. 209–219, 2006.

[3] E. C. Brown, C. T. Ragsdale, and A. E. Carter, "A grouping genetic algorithm for the multiple traveling salesperson problem," *International Journal of Information Technology and Decision Making*, vol. 6, no. 2, pp. 333–347, 2007.

[4] A. E. Carter and C. T. Ragsdale, "A new approach to solving the multiple traveling salesperson problem using genetic algorithms," *European journal of operational research*, vol. 175, no. 1, pp. 246–257, 2006.

[5] J. Ceberio, A. M. E. Irurozki, and J. Lozano, "A review on Estimation of Distribution Algorithms in Permutation-based Combinatorial Optimization Problems," *Accepted by Progress in Artificial Intelligence*, 2011.

[6] S. H. Chen, P. C. Chang, and Q. Zhang, "A self-guided genetic algorithm for flowshop scheduling problems," in *Proceedings of the Eleventh conference on Congress on Evolutionary Computation*. Institute of Electrical and Electronics Engineers Inc., The, 2009, pp. 471–478.

[7] S. H. Chen, P. C. Chang, Q. Zhang, and C. B. Wang, "A guided memetic algorithm with probabilistic models," *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 12, pp. 4753–4764, 2009.

[8] S. H. Chen and M. C. Chen, "Operators of the two-part encoding genetic algorithm in solving the multiple traveling salesmen problem," in *The 2011 Conference on Technologies and Applications of Artificial Intelligence (TAAI 2011)*, 2011.

[9] S. H. Chen, M. C. Chen, P. C. Chang, and Y. M. Chen, "Ea/g-ga for single machine scheduling problems with earliness/tardiness costs," *Entropy*, vol. 13, no. 6, pp. 1152–1169, 2011.

[10] S. H. Chen, M. C. Chen, P. C. Chang, Q. Zhang, and Y. M. Chen, "Guidelines for developing effective Estimation of Distribution Algorithms in solving single machine scheduling problems," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6441–6451, 2010.

[11] E. Falkenauer, "A hybrid grouping genetic algorithm for bin packing," *Journal of heuristics*, vol. 2, no. 1, pp. 5–30, 1996.

[12] B. Gavish and K. Srikanth, "An optimal solution method for large-scale multiple traveling salesmen problems," *Operations Research*, pp. 698–717, 1986.

[13] M. Hauschild and M. Pelikan, "An introduction and survey of estimation of distribution algorithms," *Accepted by Swarm and Evolutionary Computation*, 2011.

[14] B. Jarboui, M. Eddaly, and P. Siarry, "An estimation of distribution algorithm for minimizing the total flowtime in permutation flowshop scheduling problems," *Computers & Operations Research*, vol. 36, no. 9, pp. 2638–2646, 2009.

[15] A. H. Kasha, M. H. Kashan, and S. Karimiyan, "A particle swarm optimizer for grouping problems," *Information Sciences*, vol. 252, pp. 81–95, 2013.

[16] L. Ke, Q. Zhang, and R. Battiti, "Moea/d-aco: A multiobjective evolutionary algorithm using decomposition and ant colony," *accepted by IEEE Transactions on Systems, Man and Cybernetics, Part B*, 2013.

[17] A. Király and J. Abonyi, "A novel approach to solve multiple traveling salesmen problem by genetic algorithm," *Computational Intelligence in Engineering*, pp. 141–151, 2010.

[18] G. Laporte and Y. Nobert, "A cutting planes algorithm for the m-salesmen problem," *Journal of the Operational Research Society*, vol. 31, no. 11, pp. 1017–1023, 1980.

[19] C. J. Malmborg, "A genetic algorithm for service level based vehicle scheduling," *European Journal of Operational Research*, vol. 93, no. 1, pp. 121–134, 1996.

[20] D. Miller and J. F. Pekny, "Exact solution of large asymmetric traveling salesman problems," *Science*, vol. 251, no. 4995, p. 754, 1991.

[21] Q. Pan and R. Ruiz, "An estimation of distribution algorithm for lot-streaming flow shop problems with setup times," *Omega*, vol. 40, no. 2, pp. 166–180, 2012.

[22] Y. B. Park, "A hybrid genetic algorithm for the vehicle scheduling problem with due times and time deadlines," *International Journal of Production Economics*, vol. 73, no. 2, pp. 175–188, 2001.

[23] B. Rekiek, A. Delchambre, and H. Saleh, "Handicapped person transportation: An application of the grouping genetic algorithm," *Engineering Applications of Artificial Intelligence*, vol. 19, no. 5, pp. 511–520, 2006.

[24] R. Ruiz and C. Maroto, "A genetic algorithm for hybrid flowshops with sequence dependent setup times and machine eligibility," *European Journal of Operational Research*, vol. 169, no. 3, pp. 781–800, 2006.

[25] A. Salhi, J. A. V. Rodríguez, and Q. Zhang, "An estimation of distribution algorithm with guided mutation for a complex flow shop scheduling problem," in *Proceedings of the 9th annual conference on Genetic and evolutionary computation*. ACM, 2007, pp. 570–576.

[26] V. A. Shim, K. Tan, and C. Cheong, "A hybrid estimation of distribution algorithm with decomposition for solving the multiobjective multiple traveling salesman problem," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 42, no. 5, pp. 682–691, 2012.

[27] A. Singh and A. S. Baghel, "A new grouping genetic algorithm approach to the multiple traveling salesperson problem," *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, vol. 13, no. 1, pp. 95–101, 2009.

[28] L. Tang, J. Liu, A. Rong, and Z. Yang, "A multiple traveling salesman problem model for hot rolling scheduling in shanghai baoshan iron & steel complex," *European Journal of Operational Research*, vol. 124, no. 2, pp. 267–282, 2000.

[29] E. Vallada and R. Ruiz, "A genetic algorithm for the unrelated parallel machine scheduling problem with sequence dependent setup times," *European Journal of Operational Research*, vol. 211, no. 3, pp. 612–622, 2011.

[30] S. Y. Wang, L. Wang, M. Liu, and Y. Xu, "An effective estimation of distribution algorithm for solving the distributed permutation flowshop scheduling problem," *International Journal of Production Economics*, vol. 145, no. 1, pp. 387–396, 2013.

[31] Q. Zhang, J. Sun, and E. Tsang, "An Evolutionary Algorithm With Guided Mutation for the Maximum Clique Problem," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 2, pp. 192–200, 2005.

[32] Y. Zhang and X. Li, "Estimation of distribution algorithm for permutation flow shops with total flowtime minimization," *Computers & Industrial Engineering*, vol. 60, no. 4, pp. 706–718, 2011.

0A0-00

Quantity and Price Indicator for Technical Analysis in the Stock Market

S. H. Chen[*1] Y. H. Wang[*2] C. B. Wang[*3]

Dept. of Electronic Commerce Management, Nanhua University, Taiwan

The two things equity investors are most concerned with are: 1) stock picking - finding a good investment target, and 2) timing of purchase and sale. That said, these two things will be the main theme of this research. Although common technical indicators are being widely used by investors to determine time to enter market, these techniques have been used by too many, obtaining these information have been made relatively easy, thus making these references less valuable. In addition, past technical indicators did not take into consideration the relationship between stock prices and trading volume. Many researches have shown that using a single variable as the basis of investment is insufficient and unwise. In view of the above, this research will propose a new technical indicator that takes into consideration both stock price and trading volume at the same time, thus making it superior in terms of investment timing. This research involves seven scenarios of the price-volume relationship, converts the daily transaction data of individual stocks into daily scores, then takes the scores and transforms those into short-term, as well as long-term moving averages. The crossing of the two averages will be used to predict the trend of the stock prices in the future, thus indicating the timing of investments. For example, when short-term moving average line breaks above the long-term moving average, it indicates a buy; when short-term moving average falls below the long-term moving average, it indicates a sell. Based on the empirical test results, the performance of the strategy discussed in this research was comparable with historical records, but involves less transactions while being easy to use. We can conclude from this that the new technique can provide investors a more valuable market reference.

1. Introduction

Technical analysis of stock trends is used to help investors with timing decisions, while the most frequently used technical indicators revolve around stock prices (Li, 2010; Wu, 2010; Guo, 2008; Lin, and Li, 2006), and only a few are based on trading volumes (Cambell et al., 1992). Blume et al. (1994) indicated that trading volume can reflect new information in a timely manner, and changes in trading volume can efficiently reveal market reaction to the information. Relevant studies have proven the importance of trading volume (Gao, 2008; Lai et al., 2008). In view of the researches done by the above mentioned scholars, it can be said that both changes in stock price and changes in trading volumes have significant impact on the stock market's movement and trends, and none can be left alone.

However, when looking at the commonly used technical indicators, it can be found that regardless of whether it is a price-based or a volume-based technical indicator, the stock price or the trading volume is the sole variable that is being considered. For this reason, movements in the stock price and trading volume cannot be measured simultaneously. The first and foremost objective of this research is to develop a bivariate technical indicator that takes into account both stock price and trading volume, and this technique would assist investors in making superior timing decisions for trades. We hereby name this technical indicator the Quantity and Price Indicator (QP).

This research takes stock's daily price and trading volume, and converts those into QP index values using the pre-determined conditions. Then, by calculating the QP Index, the long term exponential moving average (EMA) and short term EMA can be constructed, and crossings of the lines would serve as buying or selling signals. In order to test the QP Index proposed in this research, empirical research will be conducted in both stock markets in the United States and Taiwan, and the result will be compared against previous literatures. Section Two contains a detailed description of the QP Indicator in explaining how it takes into consideration both stock's closing price and trading volume. Empirical result of proposed technique is compared to previous researches in Section Three, while Section Four is the conclusion of this research.

2. Research Method

Both stock's price and trading volume are crucial factors that affect investment decisions. However, technical indicators in the past are not able to consider both trading volume and price at the same time. That said, this research proposes a new Quantity and Price Indicator (QP), which requires inputs from daily closing prices and trading volumes, then translates those into a daily QP Index value. Details of the calculation procedures are outlined in Section 2.1. After obtaining the QP Index value, the long term and short term QP Index moving averages can be constructed. Based on the property and moving averages, crossings of the two moving average lines would be seen as signals for buy or sell. Section 2.2 includes a detailed description of this theory. Steps for calculating the return on investment is listed in Section 2.3.

2.1 QP Indicator

Building on the basis of the Cumulative Price and Volume Scoring System designed by Hu (2009), QP Indicator uses stock's daily price and volume relationship and defines a set of rules to evaluate each stock based on the changes of the stock's closing price and trading volume from the previous day. Also taking into consideration previous research conclusion of increased possibility of price rebound or reversal following significant volume increases, the QP Indicator's decision

Contact: Shih-Hsin Chen, Dept. of Electronic Commerce Management, Nanhua University, No. 55, Sec. 1, Nanhua Rd., Zhongkeng, Dalin Township, Chiayi County 62248, Taiwan (R.O.C.), +886-5-272-1001 ext. 56534, +886-5-242-7197, shihhsin@mail.nhu.edu.tw

rule has included the factor of trading volume increases. With this, there would be a total of seven price-volume patterns, serving as the basis for evaluating each individual stock.

QP Indicator is given a score of 3 when there is an increase to stock price while trading volume surges to 1.3 times of the 5-day moving average. The second scenario is when the stock price rises and trading volume shows an increase of 5% to 30% from the 5-day moving average, the QP Indicator is given a score of 2. The third scenario happens when the stock price moves up for two consecutive days but trading volume has fallen, the QP Indicator is given a score of 1.

The fourth rule applies when the stock price falls for two consecutive days while trading volume decreases, QP Indicator receives a score of -1. The fifth rule applies when there is downward movement accompanied by a 5% to 30% increase in trading volume to the 5-day moving average, the QP Indicator receives a score of -2. According to the sixth rule, the QP Indicator receives a score of -3 when the stock price falls and the trading volume surges to 1.3 times of the 5-day moving average. Finally, when closing price is unchanged or when none of the above scenario applies, the QP Indicator is given a score of 0. Using mathematical formulas, the seven situations on the QP Indicator proposed in this study can be expressed as below:

t : Trading day

Pt : Stock's closing price on day t

Qt : Stock's trading volume on day t

$Qma5$: Moving average on trading volume of the most recent 5 days

$ft(x)$: Score for QP Indicator on day t

$$ft(x)=\begin{cases} 3, if\ Pt>Pt-1\ and\ Qt>Qma5*1.3 \\ 2, if\ Pt>Pt-1\ and\ Qma5*1.3>Qt\ and\ Qt>Qma5*1.05 \\ 1, if\ Pt>Pt-1\ and\ Pt-1>Pt-2\ and\ Qt<Qt-1 \\ -1, if\ Pt<Pt-1\ and\ Pt-1<Pt-2\ and\ Qt<Qt-1 \\ -2, if\ Pt<Pt-1\ and\ Qma5*1.3>Qt\ and\ Qt>Qma5*1.05 \\ -3, if\ Pt<Pt-1\ and\ Qt>Qma5*1.3 \\ 0, f\ Pt=Pt-1\quad or\ otherwise \end{cases} \tag{1}$$

Explanation of Method:

1. Consider individual stock's daily change in price and volume. When today's price (Pt) is higher than that of the day before ($Pt-1$), and trading volume today (Qt) surges to 1.3 times of the 5-day moving average ($Qma5$).
2. When price today (Pt) is higher than that of the previous day ($Pt-1$), and when trading volume today (Qt) is 5% to 30% more than that of the previous day ($Qt-1$).
3. When price today (Pt) is higher than that of the previous day ($Pt-1$), and when the price on the previous day ($Pt-1$) is higher than that of the day before ($Pt-2$), while trading volume for today (Qt) is less than that of the previous day ($Qt-1$).
4. When price today (Pt) is lower than that of the previous day ($Pt-1$), and price of the previous day ($Pt-1$) is lower than that of the day before ($Pt-2$), while trading volume for today (Qt) is less than that of the previous day ($Qt-1$).
5. When the price today (Pt) is lower than that of the previous day ($Pt-1$), and when the trading volume today (Qt) is higher than that of the previous day ($Qt-1$) by 5% to 30%.
6. When the price today (Pt) is lower than that of the previous day ($Pt-1$), and when the trading volume today (Qt) surges to 1.3 time of the 5-day moving average ($Qma5$).
7. When price today (Pt) is the same as that of the previous day ($Pt-1$), or when none of the previous six scenarios applies.

2.2 Long term and short term moving averages based on QP Indicator

With the calculation method described above, each individual stock has a daily score that ranges between 3 and -3. Then based on the property of moving averages, the short term moving average ($Avgx$) and long term moving average ($Avgy$) are constructed respectively based on each stock's daily score. When the short term x-day moving average line crosses the y-day moving average line from below while the value of the x-day moving average line is lower than 0, then it signals for buy (also called the Golden Crossover). On the other hand, when the x-day moving average line crosses the the y-day moving average line from above while the value is bigger than 0.5, then it signals for sell (Death Cross). Below are the definitions and formulas for the short term and long term moving averages:

Avg : Moving average

x, y : x and y each stands for the number of days for the short term and long term moving average

ft : Value of the QP Indicator on day t

$$Avg_x = (\textstyle\sum_{i=1}^{x} f_i)/x\,, \tag{2}$$

$$i = t-x, t-x+1, t-x+2, \ldots, t-1\,,$$

Buy and sell signals from the cross up or cross down of short term and long term moving average lines as follows:

Golden Cross: $Avg_x > Avg_y$, and $Avg_x < 0$ (3)

Death Cross: $Avg_x < Avg_y$, and $Avg_x > 0.5$ (4)

2.3 Measuring return on investment

The return on investment for each trade (R) is calculated by subtracting the closing price on the purchase day (BC) from closing price on the selling day (SC), then dividing that by the closing price on the purchase day (BC), and multiply that by 100%. The calculation is explain as below:

SC : Closing price on the selling day

BC : Closing price on the purchase day

R : Sum of return on investment

Ri : Return on investment of each stock on one trade

i : Each trade

n : Number of trades

$$Ri = \frac{SCi - BCi}{BCi} * 100\,\% \quad (5)$$

Adding up the return on investment for each trade yields the sum of return on investment for each individual stock.

$$R = \sum_{i=0}^{n} Ri \quad (6)$$

3. Analysis of empirical results

To prove the validity of the technique proposed herein, the research will compare itself against various previous literatures (Chang et al., 2011; Chang, Fan & Liu, 2009; Giles, Lawrence & Tsoi, 2001; Mallick, Lee & Ong, 2008), of which all of the academics involved have proposed new methodologies to predict buy and sell points. The comparison will be made in the context of a total of six U.S. stocks while taking into consideration of three different market trends, being upward, flat and downward markets. The stocks selected for the empirical studies in the previous literatures also included U.S. and Taiwanese stocks. U.S. stocks include Apple Inc. (AAPL), The Boeing Company (BA) and Verizon Communications Inc. (VZ), which each represents for upward, sideways, downward market trends respectively.

The time interval also follows that of the previous literature. Training period for the U.S. stocks was between January 2, 2008 to December 30, 2008, while the testing period was between January 2, 2009 to June 30, 2009.

At the same time, since the new indicator utilizes the concept of moving averages, the collection of data used for purpose of comparison would have to be pushed to earlier to avoid interference to the calculation of x-day and y-day moving averages. That said, the actual testing data would have to be adjusted backward to the appropriate date; for example, for a 10-day long term moving average, testing data would have to be adjusted backward for 9 days.

3.1 Results for the training interval

This research translates each stock's daily changes in price and trading volume into daily scores. The scores ranges between 3 and -3, and then are transformed into x-day short term and y-day long term moving average lines. Through the crossing of the two moving average lines, buy and sell opportunities are being identified. Below, (x, y) will be used to represent value of a combination of short term and long term moving averages.

Based on the return on investment from the trainning results, the best days combination for Apple Inc. was (6,14), representing the buy/sell point at MA6 > MA14 which is the Golden Cross, and MA6 < MA14 is the Death Cross. The best days combinations are (6,12) for The Boeing Company, (5,10) for Verizon Communications Inc., (6,10) for AU Optronics Corporation, (6,13) for EPISTAR, and (5,9) for United Microelectronics Corporation respectively.

3.2 Empirical results for the testing interval

U.S. Stocks selected for the empirical study include Apple Inc., The Boeing Company, and Verizon Communications Inc. According to the training data, the term combination for Apple Inc. would be (6,14). Looking at the short term and long term moving average lines of Apple Inc. (see Figure 1), and the actual buy / sell points during the testing period (please refer to Figure 2), a Golden Cross appeared on January 23, 2009. On the same date, the stock would be purchased at a price of \$88.36. This holding would be sold on the Death Cross that happened on February 25, 2009 at \$96.46, resulting in a gain of 9.17%. The second purchase would be on March 6, 2009 at \$85.30, and then would be sold on April 17, 2009 at \$123.42, resulting in a 44.69% gain. The third transaction would happen on June 23, 2009 at \$134.01, and then would be sold on June 30, 2009 at \$142.43 with a gain of 6.28%.

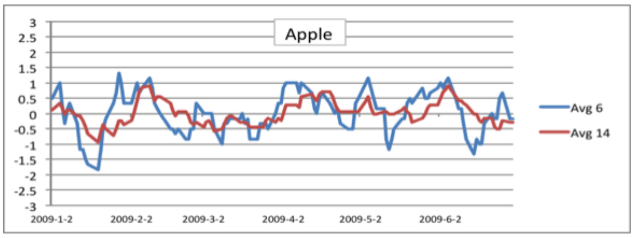

Figure 1: QP moving average lines for Apple Inc.

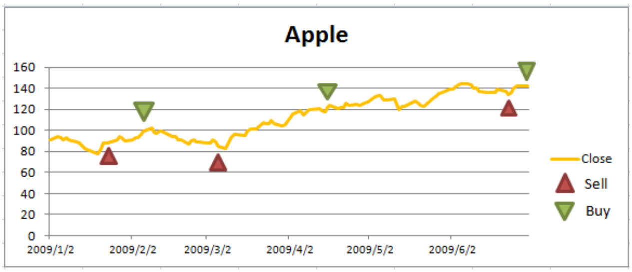

Figure 2: Buy and sell points for Apple Inc.

Table 1: Transaction records on Apple Inc.

Item	Date of Purchase	Purchase Price	Date of Sell	Selling Price	Rate of Return (%)
1	2009/1/23	88.36	2009/2/5	96.46	9.17
2	2009/3/6	85.3	2009/4/17	123.42	44.69
3	2009/6/23	134.01	2009/6/30	142.43	6.28
Total					60.14

Three trades were conducted during the empirical testing period, and the rate of return for the trades were 9.17%, 44.69% and 6.28% respectively, for a total return of 60.14%. The average rate of return per transaction would be 20.05%.

According to the training data, the term combination for The Boeing Company would be (6,12). The short term and long term moving average lines for the second stock, being The Boeing Company, as well as the actual buy/sell points during the testing period are shown in Figure 3 and 4. Because the QP Indicator's short term moving average line crossed up through the long term moving average line on January 22, 2009, a purchase for BA would be made on the same day at a price of $42.26. This purchase would then be sold on February 9, 2009 at a price of $42.80, yielding a return of 1.28%. The second transaction would happen on March 3, 2009 at a purchase price of $29.36, and would then be sold on May 6, 2009 for a return of 50.54%. The same principal for trading applies for VZ.

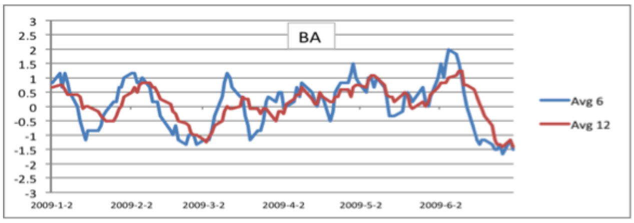

Figure 3: QP moving average lines for The Boeing Company

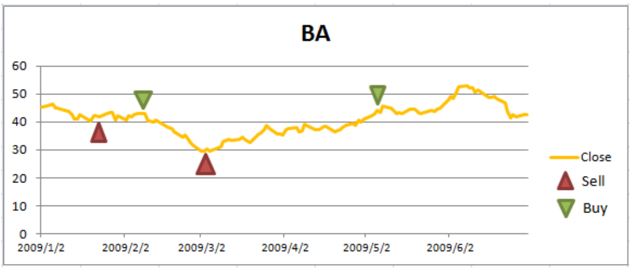

Figure 4: Buy and sell points for The Boeing Company

Table 2: Transaction records on The Boeing Company

Item	Date of Purchase	Purchase Price	Date of Sell	Selling Price	Rate of Return (%)
1	2009/1/22	42.26	2009/2/9	42.8	1.28
2	2009/3/3	29.36	2009/5/6	44.2	50.54
Total					51.82

Two trades were conducted during the empirical testing period, and the rate of return for the trades are 1.28% and 50.54% respectively, for a cumulative return of 51.82%. The average rate of return per transaction would be 25.91%.

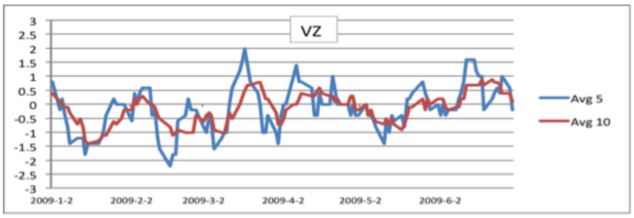

Figure 5: QP moving average lines for Verizon Communications Inc.

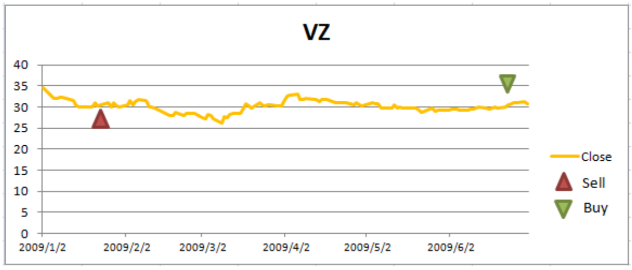

Figure 6: Buy and sell points for Verizon Communications Inc.

Table 3: Transaction records on Verizon Communications Inc.

Item	Date of Purchase	Purchase Price	Date of Sell	Selling Price	Rate of Return (%)
1	2009/1/22	30.16	2009/6/24	30.8	2.12
Total					2.12

3.3 Comparison with previous literatures

A total of six stocks were being selected as subjects for this research, three of which were U.S. stocks. The study utilized the new technical indicator that is proposed by this report.

During the testing period, twelve transactions were being conducted. The average rate of return for the six stocks was 38.02%. Chang e al. (2011) conducted a total of 54 trades with a average return of 38.22%. The table below is a comparison between the results of this research and previous literatures.

Table 4: Rate of return and number of trades for individual stocks

Literature	Stock	APPL	BA	VZ	Cumulative Average
Giles et al. (2001)	Rate of Return (%)	8.02	-20.42	13.42	0.34
	No. of trades	2	2	4	
Mallick et al. (2008)	Rate of Return (%)	10.2	15.38	12.94	12.84
	No. of trades	10	14	8	
Chang et al. (2009)	Rate of Return (%)	12.97	17.5	**27.72**	19.40
	No. of trades	23	20	11	
Chang et al. (2011)	Rate of Return (%)	**61.28**	38.03	15.36	**38.22**
	No. of trades	13	11	4	
This research	Rate of Return (%)	60.14	**51.82**	2.12	38.02
	No. of trades	3	2	1	

Out of all the stocks, Apple and Boeing were the best performing ones, with each of them ending up with a return of over 50 percent. The average rate of return of all the stocks was 38.02%, which was slightly lower than the 38.22% in previous literatures. However, the number of transactions for the previous literature was fifty two times, which is a lot higher than that the twenty eight trades for this research. When taken into account transaction costs, return from this research would be higher than that of the literature.

From the empirical test, it could be proven that although return from this research is slightly lower than that of the literature, but the average rate of return per trade would be significantly higher. That said, if transaction cost is a constraining factor that needs to be taken into consideration, this research results in a lower fee. With this in mind, the QP Indicator that is proposed in this study would be of value to investors.

4. Conclusion and recommendation

This research proposes a new QP Indicator as a technical indicator that not only takes into consideration changes in stock prices, but also emphasizes on the value of monitoring trading volume. This new method employes the property of moving averages to calculate short term and long term moving average lines, and utilizes the relationship between these lines to identify the best timing for trades. Through comparison with previous literatures, this new method is seen to be able to generate gains for stocks in upward, sideways, and downward trends. Although the cumulative return on investment is less than 0.5% shy of that from literature, the number of trades is limited to twelve times, which is less than the fifty four times required to generate the higher return from the previous study. This decreased number of transactions avoids the case of frequent trading signals, as well as the transaction costs (or transaction taxes) that are triggered by the excessive trading that would erode on gains. Taking the related costs into consideration, the performance of the method proposed by this research would exceed the return of previous studies. Other than that, this new indicator would be applicable to the U.S. markets.

The advantage of the QP Indicator would be its ease of understanding and execution. Only simple calculations are required to identify superior buying and selling opportunities. Also, different combinations of short term and long term moving averages would allow investors the flexibility to choose between suitable parameters for short term, medium term, and long term investments.

References

[Beaver 1970] Beaver, W., Kettler, P., & Scholes, M., *The association between market determined and accounting determined risk measures.*, In Psychological Review, 35, 654-682, 1970.

[Blume 1994] Blume, L., Easley, D., & O'Hara, M., *Market statistics and technical analysis: The role of volume.*, Journal of Finance, 49(1), 153-181, 1994.

[Bondt 1985] Bondt, W. F. M. D., & Thaler, R., *Does the stock market overreact?*, Journal of Finance, 40(3), 793-805, 1985.

[Chang 2009] Chang, P. C., Fan, C. Y., & Liu, C. H., *Integrating a piecewise linear representation method and a neural network model for stock trading points prediction.*, Applications and Reviews, 39(1), 80-92, 2009.

[Chang 2011] Chang, P. C., Liao, T. W., Lin, J. J., & Fan, C. Y., *A dynamic threshold decision system for stock trading signal detection.*, Applied Soft Computing, 11(5), 3998-4010, 2011.

[Conrad 1988] Conrad, J., & Kaul, G., *Time-Variation in Expected Returns.*, Journal of Business, 61, 409-425, 1988.

[Edwards 1992] Edwards, R. D., & Magee, J., *Technical analysis of stock trends.*, Boston: J. Magee Inc., 1992.

[Gao 2008] Gao, S. X., *Price-Volume Relationship: Is Volume a Leading Indicator of Price Discovery*., Feng Chia University Department of Finance, 2008.

[Gervais 2001] Gervais, S., Kaniel, R., & Mingelgrin, D. H., *The High-Volume Return Premium.*, The Journal of Finance, 56(3), 877-919, 2001.

[Giles 2001] Giles, C. L., Lawrence, S., & Tsoi, A. C., *Noisy time series prediction using recurrent neural networks and grammatical Inference.*, Machine Learning, 44(1-2), 161-183, 2001.

[Guo] Kuo, H. Y., *The Empirical Study of Technical Analysis in China Stock Market*., Unpublished master's thesis, Ling Tung University of Science and Technology, 2008.

[Hu 2009] Hu, L. Y., *The 100 methods to invest on stock market, Economic Daily News*., Taipei, 2009.

[Lai 2008] Lai, Y. C., Chang, C. M., Liao, M. L., Huang, C. L., *The Study on the Price-Volume Relationship of Taiwan Stock Index, Taiwan Stock Index Futures, and MSCI Taiwan Stock Index Futures: The Application of GARCH Model.*, Journal of Chinese Economic Research, 6(2), 18-34, 2008.

[Li 2010] Li, G., *The significance of technical analysis in the stock market.*, The management & technology of the small and medium enterprises, 2010(8A), 73-74, 2010.

[Lin 2006] Lin, J. X. and Li, S. H., *Profitability of technical indicators Quote decline in non-linear relationship.*, Taiwan Academy of Management Journal, 2006.

[Mallick 2008] Mallick, D., Lee, V. C., & Ong, Y. S., *An empirical study of Genetic Programming generated trading rules in computerized stock trading service system.*, Service Systems and Service Management, 2008 International Conference on 1-6, 2008.

[McInish 1991] McInish, T. H., & Wood, R. A., *Hourly Returns, Volume, Trade Size, and Number of Trades.*, Journal of Financial Research, 14(4), 303-315, 1991.

[Pruitt 1998] Pruitt, S. W., & White, R. E., *The CRISMA trading system：Who says technical analysis can't beat the market?*, The Journal of Portfolio Management, 14, 55-58, 1998.

[Wu 2001] Wu, T. A., & Gau, Y. F., *Relation between volatility and volume.*, Unpublished master's thesis, National Chi Nan University, 2001.

.

國科會補助計畫衍生研發成果推廣資料表

日期:2013/10/17

<table>
<tr><td rowspan="3">國科會補助計畫</td><td>計畫名稱: 估計機率分佈演算法求解群內最佳化問題之研究</td></tr>
<tr><td>計畫主持人: 陳世興</td></tr>
<tr><td>計畫編號: 101-2221-E-230-026- 學門領域: 人工智慧與仿生計算</td></tr>
<tr><td colspan="2">無研發成果推廣資料</td></tr>
</table>

101 年度專題研究計畫研究成果彙整表

計畫主持人：陳世興	計畫編號：101-2221-E-230-026-
計畫名稱：估計機率分佈演算法求解群內最佳化問題之研究	

成果項目			量化：實際已達成數（被接受或已發表）	量化：預期總達成數(含實際已達成數)	量化：本計畫實際貢獻百分比	單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	4	4	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	1	100%	篇	We will submit a paper to the IJPR.
		研究報告/技術報告	1	1	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估 ■達成目標 □未達成目標（請說明，以 100 字為限） □實驗失敗 □因故實驗中斷 □其他原因 說明：
2. 研究成果在學術期刊發表或申請專利等情形： 論文：□已發表 □未發表之文稿 ■撰寫中 □無 專利：□已獲得 ■申請中 □無 技轉：□已技轉 □洽談中 ■無 其他：（以 100 字為限）
3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限） Because there is quite a few research of EDA studied the in-group optimization problems, such as the multiple traveling salesmen problem (mTSP) and parallel machine scheduling problem (PMSP). This research proposed an algorithm deal by using the Self-Guided GA together with the Minimum Loading Assignment rule (MLA) to tackle the mTSP. The solution space of the proposed method would be only n!. We compare the proposed algorithm against the best direct encoding technique, two-part encoding genetic algorithm (TPGA) and its solution space is n!*C(n-1, m-1). From the experimental results, the proposed algorithm outperforms the TPGA significantly. It shows a good direction for the EDAs researchers who like to solve the in-group optimization problems.